

文章编号: 1673-3193(2024)06-0737-07

基于注意力机制和小波变换的图像隐写方法

李佩佩^{1,2}, 陈燕^{1,2}, 王康谊³, 张权^{1,2}

(1. 中北大学 信息与通信工程学院, 山西 太原 030051;

2. 中北大学 生物医学成像与影像大数据山西省重点实验室, 山西 太原 030051;

3. 中北大学 创新与创业学院, 山西 太原 030051)

摘要: 为了更好地平衡隐写容量和不可感知性之间的关系, 本文提出了一种基于SteganoGAN的优化方案。首先, 将SteganoGAN隐写网络和提取网络进行加深, 以增强模型的复杂度和学习能力; 其次, 为了实现更为隐蔽的信息嵌入, 在隐写网络部分引入离散小波变换(Discrete Wavelet Transform, DWT)和逆离散小波变换(Inverse Discrete Wavelet Transform, IDWT)模块, 这使得秘密信息能够被有效地嵌入到图像的小波域中; 最后, 在隐写网络、提取网络中融入了一种改进的通道-空间注意力模块(Improved Channel and Spatial Attention Module, ICAM-SAM), 促使模型能够聚焦于图像中的高隐蔽性区域, 实现更为精准的信息隐藏。实验结果表明: 改进后的模型在提取准确率上提高了0.84个百分点, 表明其隐写和提取过程更加精确。此外, 每像素嵌入率(Reed-Solomon Bits-Per-Pixel, RS-BPP)提高了1.71%, 这表明改进后的模型在相同大小的图像中可以隐藏更多的信息。同时, 峰值信噪比(Peak Signal to Noise Ratio, PSNR)提高了12.53%、结构相似性(Structural Similarity Index, SSIM)提高了5.14%, 这表明嵌入的信息对原始图像的影响更小, 改进后的模型具有更高的图像质量。综合结果表明, 改进后的模型具有更好的不可感知性和较大的隐写容量。

关键词: 图像隐写; 生成对抗网络; 离散小波变换; 逆离散小波变换; 改进的通道-空间注意力模块

中图分类号: TP309.2 **文献标识码:** A **doi:** 10.3969/j.issn.1673-3193.2024.06.002

引用格式: 李佩佩, 陈燕, 王康谊, 等. 基于注意力机制和小波变换的图像隐写方法[J]. 中北大学学报(自然科学版), 2024, 45(6): 737-743.

LI Peipei, CHEN Yan, WANG Kangyi, et al. Image steganography method based on attention mechanism and wavelet transform[J]. Journal of North University of China(Natural Science Edition), 2024, 45(6): 737-743.

Image Steganography Method Based on Attention Mechanism and Wavelet Transform

LI Peipei^{1,2}, CHEN Yan^{1,2}, WANG Kangyi³, ZHANG Quan^{1,2}

(1. School of Information and Communication Engineering, North University of China, Taiyuan 030051, China;

2. Shaxi Provincial Key Laboratory of Biomedical Imaging Big Data, North University of China, Taiyuan 030051, China;

3. School of Innovation and Entrepreneurship, North University of China, Taiyuan 030051, China)

Abstract: In order to better balance the relationship between steganographic capacity and imperceptibility, this paper proposed an optimization scheme based on SteganoGAN. Firstly, the steganography network and extraction network of SteganoGAN were deepened to enhance the complexity and learning capability of the model; Secondly, in order to achieve more covert information embedding, discrete wavelet transform

收稿日期: 2024-03-27

基金项目: 山西省基础研究计划项目(202103021224204)

作者简介: 李佩佩(1999-), 女, 硕士生, 主要从事图像隐写方面的研究。

通信作者: 王康谊(1967-), 女, 教授, 博士, 主要从事电子技术及测控方向的研究。E-mail: wangkangyi@nuc.edu.cn。

(DWT) and inverse discrete wavelet transform (IDWT) modules were introduced in the steganography network, which enabled secret information to be effectively embedded in the wavelet domain of the image; Finally, an improved channel and spatial attention module (ICAM-SAM) was incorporated into the steganography and extraction networks, enabling the model to focus on highly covert areas in the image and achieve more accurate information hiding. Experimental results show that the accuracy of the improved model is increased by 0.84 percentage point, indicating that its steganography and extraction processes are more accurate. In addition, the reed-solomon bits per pixel (RS-BPP) is increased by 1.71%, indicating that the improved model can hide more information in images of the same size. Meanwhile, the peak signal to noise ratio (PSNR) is increased by 12.53% and the structural similarity index (SSIM) is increased by 5.14%, indicating that the embedded information has a smaller impact on the original image and the improved model has higher image quality. In summary, these results indicate that the improved model has better imperceptibility and larger steganography capacity.

Key words: image steganography; generating adversarial networks; discrete wavelet transform; inverse discrete wavelet transform; improved channel and spatial attention module

0 引言

隐写术^[1]是一种在图像中嵌入秘密信息的技术,其目的是将秘密信息以一种难以察觉的方式隐藏在载体图像中,它不仅保证了秘密信息本身的安全,还提升了信息传输过程的安全性。与密码学相比,它更加注重隐藏通信过程本身,隐写术可以保证网络通信过程中数据的安全性,因此成为近年来信息安全领域的热门研究方向之一^[2]。

随着人工智能和深度学习的发展,各种深度学习模型已被广泛应用于图像隐写,分析现有文献,大多数学者的方法涉及生成对抗网络(Generative Adversarial Network, GAN)、小波变换和注意力机制等。GAN是Goodfellow等^[3]在2014年引入的一种深度卷积神经网络。毕新亮等^[4]使用Star GAN的判别器作为特征提取器,将提取的特征映射为秘密消息,直接构造了图像和秘密消息的映射关系。Martin等^[5]使用GAN来增强空间域隐写方法的性能,通过最小化图像修改来嵌入秘密信息,成功规避了许多先进深度学习隐写分析架构的检测。Zhang等^[6]提出了一种新的模型ISGAN,通过将秘密图像仅隐藏在载体图像的Y通道中来提高不可见性。Zhu等^[7]提出了一个名为HiDDeN的新模型,该模型引入了噪声层的概念,它可以提高模型的鲁棒性。这些研究方法虽然在一定程度上实现了图像隐写,但是在更高容量和更强不可见性方面仍面临着新的挑战。

在像素域隐藏图像容易导致纹理复制伪影和颜色失真。与像素域相比,频域尤其是高频域更适合

于图像的隐藏。Jing等^[8]使用小波变换(DWT)^[9]将图像分成低频和高频小波子带,然后再进入可逆网络,使网络能够更好地将秘密信息融合到载体图像中。Guan等^[10]也采用了将秘密信息隐藏在高频域的方式,与文献[8]不同的是,Guan等^[10]提出的网络大大提高了隐写容量,它可以在一张彩色图像中同时隐藏两张与载体图像大小相同的图片。

计算机视觉中常用的注意机制主要有两种:通道注意机制和空间注意机制。它们聚集所有位置的相关特征,并通过使用不同的聚集策略、转换和强化函数来强化原始特征。近年来,越来越多的注意力机制被提出,如SE^[11]、ECA-Net^[12]、CBAM^[13]等。Tan等^[14]设计了一个特定的通道注意力模块,利用通道之间的相互依赖动态地调整图像的深度表征,有利于提高隐写图像的质量和消息提取的准确性。Peng等^[15]提出了一种基于双重注意机制的文本隐写分析方法,可以动态地融合语义和统计特征,在3种不同的文本隐写分析场景中超越了当前较先进的技术。受此启发,本文将双通道注意力机制引入隐写网络和提取网络中,用来提高本文模型的隐写效果。

综上,本文在SteganoGAN^[16]的基础上提出了一个改进的图像隐写方法,主要贡献为:

- 1) 将隐写网络和提取网络分别由3层加深到5层,具体分别增加了式(10)中的步骤e、f和式(2)中的步骤d、e。

- 2) 在隐写网络中引入DWT和IDWT,实现了将秘密信息隐藏在小波域,提高了隐写容量。

- 3) 在隐写网络和提取网络中分别引入

ICAM-SAM, 将嵌入修改集中在纹理复杂的区域, 以改善生成式隐写算法安全性较低的问题。

1 本文方法

1.1 整体网络结构

SteganoGAN 模型包括隐写网络、提取网络和评估网络三部分, 改进后的模型结构如图 1 所示。

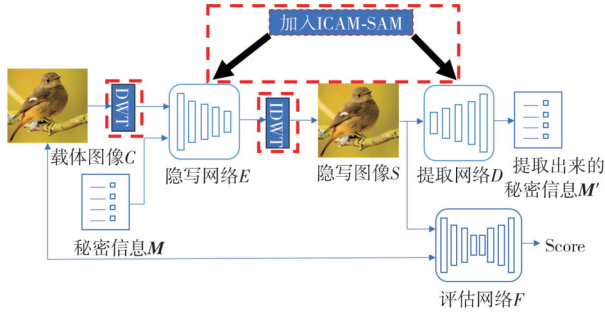


图 1 改进后的模型整体结构

Fig. 1 The overall structure of the improved model

1.1.1 隐写网络

隐写网络 E 的输入是 $3 \times W \times H$ 的载体图像和秘密信息 ($M \in \{0, 1\}^{Depth \times W \times H}$), 输出为隐写图像。其中, M 是一个形状为 $Depth \times W \times H$ 的二进制数据张量, $Depth$ 是试图在载体图像的每个像素中隐藏的位数, $W \times H$ 表示载体图像的大小。隐写网络采用了密集连接网络的结构, 后面每一层的输入为前面所有层的输出的通道拼接, 这样可以将早期模块生成的特征映射连接到后面模块生成的特征映射。隐写网络由 6 个卷积层构成, 每个卷积核的尺寸为 3×3 , 每一层的激活函数为 LeakyReLU, 最后一层不含激活函数和 BN。整个过程如式(1)所示, 其中, Conv 包括卷积、LeakyReLU 和 BN, Cat 为通道拼接。

$$\begin{aligned} a &= Conv_{3 \rightarrow 32}(C), \\ b &= Conv_{32 + Depth \rightarrow 32}(Cat(a, M)), \\ c &= Conv_{64 + Depth \rightarrow 32}(Cat(a, b, M)), \\ d &= Conv_{96 + Depth \rightarrow 32}(Cat(a, b, c, M)), \\ e &= Conv_{128 + Depth \rightarrow 32}(Cat(a, b, c, d, M)), \\ f &= Conv_{160 + Depth \rightarrow 3}(Cat(a, b, c, d, e, M)), \\ E(C, M) &= C + f. \end{aligned} \quad (1)$$

1.1.2 提取网络

提取网络 D 的输入为隐写图像 S , 输出为重建的秘密信息 M' , 它的结构与隐写网络相似, 提取网络整个过程可以表示如下

$$a = Conv_{3 \rightarrow 32}(S),$$

$$\begin{aligned} b &= Conv_{32 \rightarrow 32}(a), \\ c &= Conv_{64 \rightarrow 32}(Cat(a, b)), \\ d &= Conv_{96 \rightarrow 32}(Cat(a, b, c)), \\ e &= Conv_{128 \rightarrow 32}(Cat(a, b, c, d)), \\ D(S) &= Conv_{160 \rightarrow Depth}(Cat(a, b, c, d, e)). \end{aligned} \quad (2)$$

1.1.3 评估网络

为了对隐写网络的性能提供反馈并生成更逼真的隐写图像, 引入了评估网络 F , 它可以用来区分载体图像 C 和隐写图像 S , 评估网络整个过程可以表示为

$$\begin{aligned} a &= Conv_{32 \rightarrow 32}(Conv_{32 \rightarrow 32}(Conv_{3 \rightarrow 32}(S))), \\ C(S) &= Mean(Conv_{32 \rightarrow 1}(a)). \end{aligned} \quad (3)$$

1.2 离散小波变换 (DWT) 和逆离散小波变换 (IDWT)

DWT 可将图像转换为频域表示, 使得嵌入操作在频域中进行, 增加了嵌入容量。本文的改进之一是将 DWT 和 IDWT 引入到隐写网络中。首先将载体图像经过 DWT, 分解出 A 、 H 、 V 、 D (见式(5)) 4 个子带图像, 接下来将分量 A 、 H 、 V 、 D 进行通道拼接, 输入隐写网络, 然后再对输出进行 IDWT, 得到隐写网络的输出, 此过程见式(4)。其中, $Inputs$ 和 $Outputs$ 分别表示输入和输出, $DWT(\cdot)$ 表示离散小波变换, $IDWT(\cdot)$ 表示逆离散小波变换, $Cat(\cdot)$ 表示通道拼接, $E(\cdot)$ 表示隐写网络。由于 Haar 小波具有简单有效、双向对称的特点, 其应用不会影响网络的端到端训练。

$$A, H, V, D = DWT(Inputs),$$

$$Outputs = IDWT(E(Cat(A, H, V, D))). \quad (4)$$

具体地, 通过使用 4 个卷积滤波器来实现二维的 DWT, 将输入图像 I 分解为 4 个子带图像, 即 I_{LL} (平均分量 A)、 I_{HL} (水平分量 H)、 I_{LH} (垂直分量 V) 和 I_{HH} (对角分量 D), 使用的滤波器包括 $k_{LL} = [1 \ 1; 1 \ 1]^T$, $k_{HH} = [1 \ -1; -1 \ 1]^T$, $k_{HL} = [-1 \ 1; -1 \ 1]^T$ 和 $k_{LH} = [-1 \ -1; 1 \ 1]^T$ 。

$$\begin{aligned} I_{LL}(i, j) &= I(2i - 1, 2j - 1) + I(2i - 1, 2j) + \\ &I(2i, 2j - 1) + I(2i, 2j), \end{aligned}$$

$$\begin{aligned} I_{HL}(i, j) &= -I(2i - 1, 2j - 1) + I(2i - 1, 2j) - \\ &I(2i, 2j - 1) + I(2i, 2j), \end{aligned}$$

$$\begin{aligned} I_{LH}(i, j) &= -I(2i - 1, 2j - 1) - I(2i - 1, 2j) + \\ &I(2i, 2j - 1) + I(2i, 2j), \end{aligned}$$

$$\begin{aligned} I_{HH}(i, j) &= I(2i - 1, 2j - 1) - I(2i - 1, 2j) - \\ &I(2i, 2j - 1) + I(2i, 2j). \end{aligned} \quad (5)$$

与像素域相比,频域尤其是高频域更适合信息的隐藏。DWT可以将图像从像素域转到频域,然后再输入隐写网络,从而使网络能够更好地将秘密信息融合到载体图像中。

1.3 改进的通道-空间注意力模块 (ICAM-SAM)

Li等^[17]在网络的不同层引入了注意力机制,提高了隐写图像的质量和模型对隐写变化的敏感性。在此基础上,本文尝试在隐写网络和提取网络每一层卷积后加入ICAM-SAM,注意力机制可以使模型将嵌入的重点集中在图像中相对不容易引起怀疑的区域。将改进的通道注意力模块(Improved Channel Attention Module, ICAM)和改进的空间注意力模块(Improved Spatial Attention Module, ISAM)结合使用,共同提升隐写模型的性能,ICAM、ISAM的具体结构如图2、图3所示,整个ICAM-SAM的结构如图4所示。

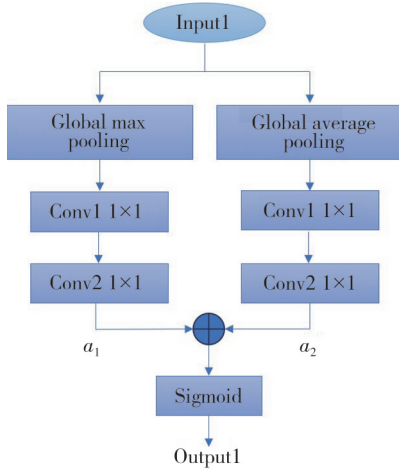


图2 ICAM结构图

Fig. 2 The structural diagram of ICAM

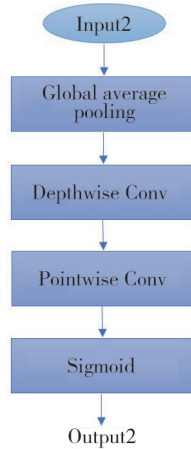


图3 ISAM结构图

Fig. 3 The structural diagram of ISAM

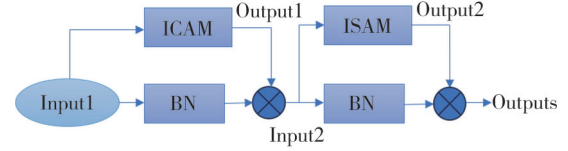


图4 ICAM-SAM结构图

Fig. 4 The structural diagram of ICAM-SAM

式(6)中, $Input1$ 和 $Input2$ 分别表示图2和图3的输入, $Output1$ 和 $Output2$ 分别表示图2和图3的输出, $Outputs$ 为ICAM-SAM的输出(见图4), a_1 和 a_2 分别为图2中左侧和右侧的中间变量, $\sigma(\cdot)$ 表示Sigmoid激活函数, $f^{1 \times 1}(\cdot)$ 表示 1×1 的卷积, $Global AvgPool(\cdot)$ 表示全局平均池化, $Global MaxPool(\cdot)$ 表示全局最大池化, $PW(\cdot)$ 表示逐点卷积(Pointwise Conv), $DW(\cdot)$ 表示深度卷积(Depthwise Conv), $BN(\cdot)$ 表示批处理归一化。

$$a_1 = f^{1 \times 1} \left(f^{1 \times 1} (Global MaxPool(Input1)) \right),$$

$$a_2 = f^{1 \times 1} \left(f^{1 \times 1} (Global AvgPool(Input1)) \right),$$

$$Output1 = \sigma(a_1 + a_2),$$

$$Output2 =$$

$$\sigma \left(PW \left(DW (Global AvgPool(Input2)) \right) \right),$$

$$Input2 = BN(Input1) \otimes Output1,$$

$$Outputs = BN(Input2) \otimes Output2. \quad (6)$$

2 实验结果及分析

2.1 实验设置

实验环境为 Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40 GHz, GPU 为 NVIDIA GeForce GTX 1660, 内存 32 GB, 采用 Python3.8 和 Pytorch1.13.0。实验中,共使用了两个数据集,在实验 2.3.1 和 2.3.2 中使用的是 Div2k^[18] 数据集,在实验 2.3.3 中使用的是 COCO^[19] 数据集。在每次迭代中,将每张载体图像与随机生成的序列 $Depth \times W \times H$ 组成的数据张量 M 进行匹配,该序列从伯努利分布 $M \sim Ber(0.5)$ 中采样。此外,在预处理中使用标准的数据增强过程,包括水平翻转、随机裁剪等等。使用学习率为 1×10^{-4} 的 Adam 优化器,将梯度范数归一化为 0.25,训练总次数 epoch 为 32。

2.2 评价指标

采用准确率(ACC) R_{ACC} 、每像素嵌入率(RS-

BPP) $R_{RS-BPP}^{[13]}$ 、峰值信噪比(PSNR) R_{PSN} 、结构相似性(SSIM) R_{SSIM} 来客观评价算法效果。

准确率是衡量隐写算法的重要准则之一,较高的准确率意味着预测的信息 M' 更接近原始信息 M 。对于模型容量,使用 RS-BPP 表示图像中可靠传输的平均比特数

$$R_{RS-BPP}(k, p) = k \times (2p - 1), \quad (7)$$

式中: k 表示消息长度; p 表示给定模型能够正确解码嵌入消息的一位的概率。

PSNR 则用来衡量载体图像与隐写图像之间的峰值误差。

$$E_{PSNR} = 10 \cdot \lg \frac{P_{\max}^2}{E_{MS}}, \quad (8)$$

表 1 本文模型与 SteganoGAN 模型的对比结果

Tab. 1 Comparison results between own model and SteganoGAN model

数据集	Depth	Ours			SteganoGAN				
		$R_{ACC}/\%$	R_{RS-BPP}/bpp	R_{PSN}/dB	R_{SSIM}	$R_{ACC}/\%$	R_{RS-BPP}/bpp	R_{PSN}/dB	R_{SSIM}
Div2k	1	99.75	0.995 1	44.595 9	0.981 8	98.91	0.978 3	39.628 5	0.933 8
	2	98.53	1.941 2	38.524 8	0.930 1	97.88	1.915 4	38.083 4	0.889 9
	3	94.38	2.662 8	37.241 3	0.864 3	91.52	2.491 2	36.309 2	0.840 1

由表 1 可知,在同一个 $Depth$ 中,与 SteganoGAN 模型相比,本文模型几乎在所有指标上都有较好的表现。随着 $Depth$ 的增加,两个模型的提取准确率、PSNR 和 SSIM 都呈现下降的趋势,而嵌入容量呈现上升的趋势。对实验结果进行分析,可以发现本文模型的 4 个评价指标在数值上都有所提高,但在准确性方面的改善并不明显。

为了更好地验证本文模型隐写图像的质量,对载体图像和隐写图像的直方图进行了分析比较。本文模型的载体图像、隐写图像和直方图如图 5 所示。

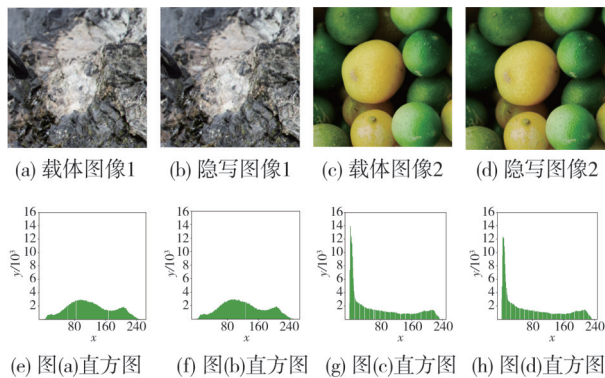


图 5 载体图像、隐写图像及其直方图

Fig. 5 Cover images, steganographic images and their histograms

经过对比分析,发现载体图像与隐写图像的直方图非常相似,只有一些微小的差异,这说明

式中: E_{MS} 为均方误差(Mean Square Error, MSE); P_{\max} 为载体图像或隐写图像的最大像素值。

SSIM 表示图像间的结构相似性。

$$R_{SSIM}(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)}. \quad (9)$$

2.3 实验结果

2.3.1 本文模型与 SteganoGAN 模型的对比实验

实验中,使用 Div2k 数据集训练网络模型。 $Depth$ 是试图在载体图像的每个像素中隐藏的位数,实验使用 3 个不同的数据深度 1, 2, 3 来训练本文的模型,实验结果如表 1 所示。

本文模型能够在不显著影响载体图像质量的情况下进行信息隐藏。

2.3.2 消融实验

为了研究 DWT、IDWT 和 ICAM-SAM 的有效性,设计了一系列消融实验,使用 Div2k 数据集,数据深度 $Depth=1$ 时训练各个模型,使用 ACC、RS-BPP、PSNR 和 SSIM 作为评价指标,实验结果如表 2 所示。

表 2 消融实验结果

Tab. 2 Results of ablation experiment

DWT、IDWT	ICAM-SAM	$R_{ACC}/\%$	R_{RS-BPP}/bpp	R_{PSN}/dB	R_{SSIM}
×	×	98.91	0.978 3	39.628 5	0.933 8
✓	×	99.66	0.993 2	43.379 9	0.971 6
×	✓	99.70	0.993 0	44.095 2	0.970 4
✓	✓	99.75	0.995 1	44.595 9	0.981 8

由表 2 可知,在隐写网络前后分别引入 DWT 和 IDWT 后,模型的 ACC、RS-BPP、PSNR 和 SSIM 分别提升了 0.75 百分点, 1.52%, 9.47% 和 4.05%; 在隐写网络和提取网络中引入 ICAM-SAM 后,模型的 4 个评价指标分别提升了 0.79 百分点, 1.50%, 11.27% 和 3.92%; 将 DWT、IDWT 和 ICAM-SAM 同时引入模型后,模型整体的评价指标分别提高了 0.84 百分点, 1.71%, 12.53% 和 5.14%。分析实验结果可以发现,

DWT、IDWT 和 ICAM-SAM 的引入对模型的性能提升具有显著影响。具体而言, DWT 和 IDWT 的使用有助于提高嵌入容量和图像质量, 而 ICAM-SAM 的引入则有助于提高准确率和不可感知性。这些结果证明了 DWT、IDWT 和 ICAM-SAM 的有效性。

为了进一步验证隐写图像的隐藏效果和不可感知性, 对载体图像和隐写图像的残差图像进行了分析, 并对残差图像进行了放大处理, 结果如图 6 所示。其中, 第一行是 SteganoGAN 模型的载体图像、隐写图像和残差图像, 第二行是本文模型的载体图像、隐写图像和残差图像。由图 6 可知: SteganoGAN 模型得到的隐写图像虽然不存在明显的颜色失真, 但是当残差图像放大到 10 倍、20 倍时, 可以隐约地发现载体图像的轮廓; 本文模型的残差图像放大 20 倍后, 几乎无法发现载体图像的轮廓, 故本文模型在隐写任务中表现出了良好的隐藏性能。

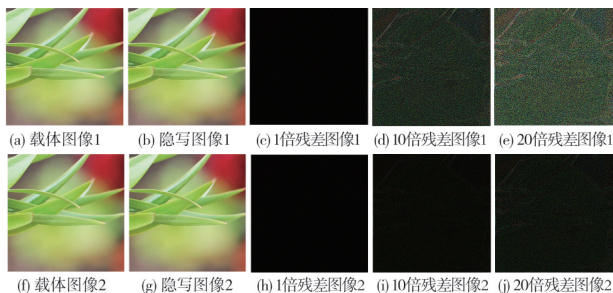


图 6 本文模型和 SteganoGAN 模型的残差图像

Fig. 6 Residual image of our model and steganoGAN model

2.3.3 不同模型的对比实验

使用 COCO 数据集, 数据深度 $Depth=1$ 时训练了各模型。在 COCO 数据集中选取 10 000 张作为训练集, 从剩余的图像中选取 1 000 张作为测试集, 实验结果如表 3 所示。

由表 3 可知: 对于 ACC 指标, HiDDeN 模型的准确率较低, 而 RoSteALS、FNNS 和本文模型的准确率都超过 99%, 说明这 3 个模型能够成功地从隐写图像中学习到隐藏信息的特征, 并将其提取出来; 对于 RS-BPP 指标, HiDDeN 和 RoSteALS 模型的值较低, 而 FNNS 和本文模型的值接近 1, 说明这两个模型在保持图像结构的同时能够嵌入更多的隐写信息; 对于 PSNR 指标, 本文模型的结果 41.977 7 dB 远高于其他模型, 说明本文模型在嵌入隐写信息的同时, 能够保持图像的高质量; 对于 SSIM 指标, 本文模型的结果为

0.980 0, 说明本文模型在保持图像的结构相似性方面表现出色。分析实验结果可以发现, 虽然本文模型的提取准确率不是最高的, 但是本文模型的嵌入容量和不可感知性达到了较高的数值, 本文模型的有效性得到了进一步验证。

表 3 不同模型的评价指标对比

Tab. 3 Comparison of evaluation indicators for different models

模型 (Depth=1)	$R_{ACC}/\%$	R_{RS-BPP}/bpp	R_{PSN}/dB	R_{SSIM}
HiDDeN	57.90	0.1579	33.1395	0.9318
RoSteALS	99.00	0.7853	28.9011	0.9647
FNNS	99.49	0.9800	30.0313	0.8436
本文算法	99.14	0.9828	41.9777	0.9800

3 结 论

为了同时提高图像隐写算法的不可感知性和容量, 本文提出了一种基于 SteganoGAN 改进的模型。本文在隐写网络前后分别引入 DWT 和 IDWT, 尝试将秘密信息隐藏在小波域, 而不是像素域; 在隐写网络和提取网络中引入了 ICAM-SAM, 以动态地优化有效负载分布。实验结果表明, 本文算法可以有效地提高隐写算法的不可感知性和容量。在今后的工作中, 将考虑使用隐写分析器替代评估器, 可能会进一步增强模型的鲁棒性。

参考文献:

- [1] SUBRAMANIAN N, ELHARROUSS O, AL-MAADEED S, et al. Image steganography: A review of the recent advances [J]. IEEE Access, 2021, 9: 23409-23423.
- [2] 陈君夫, 付章杰, 张卫明, 等. 基于深度学习的图像隐写分析综述 [J]. 软件学报, 2021, 32(2): 551-578.
CHEN Junfu, FU Zhangjie, ZHANG Weiming, et al. Review of image steganalysis based on deep learning [J]. Journal of Software, 2021, 32(2): 551-578. (in Chinese)
- [3] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Advances in Neural Information Processing Systems, 2014, 27: 2672-2680.
- [4] 毕新亮, 杨晓元, 刘文超, 等. 载体选择型图像隐写算法研究 [J]. 计算机应用研究, 2021, 38(8): 2465-2468.
BI Xinliang, YANG Xiaoyuan, LIU Wenchao, et al. Research on cover selection image steganography algo-

- rithm[J]. *Application Research of Computers*, 2021, 38(8): 2465-2468. (in Chinese)
- [5] MARTÍN A, HERNÁNDEZ A, ALAZAB M, et al. Evolving generative adversarial networks to improve image steganography[J]. *Expert Systems with Applications*, 2023, 222: 119841-119855.
- [6] ZHANG R, DONG S, LIU J. Invisible steganography via generative adversarial networks[J]. *Multimedia Tools and Applications*, 2019, 78(7): 8559-8575.
- [7] ZHU J, KAPLAN R, JOHNSON J, et al. Hidden: hiding data with deep networks [C]//*Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 657-672.
- [8] JING J, DENG X, XU M, et al. HiNet: deep image hiding by invertible network [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 4733-4742.
- [9] WU P C, CHEN L G. An efficient architecture for two-dimensional discrete wavelet transform[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(4): 536-545.
- [10] GUAN Z, JING J, DENG X, et al. DeepMIH: Deep invertible network for multiple image hiding[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(1): 372-390.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132-7141.
- [12] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 11534-11542.
- [13] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C]//*Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3-19.
- [14] TAN J, LIAO X, LIU J, et al. Channel attention image steganography with generative adversarial networks [J]. *IEEE Transactions on Network Science and Engineering*, 2021, 9(2): 888-903.
- [15] PENG W, LI S, QIAN Z, et al. Text steganalysis based on hierarchical supervised learning and dual attention mechanism [J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2023, 31: 3513-3526.
- [16] ZHANG K A, CUESTA-INFANTE A, XU L, et al. SteganoGAN: high capacity image steganography with GANs [DB/OL]. (2019-01-30) [2024-03-27]. <http://arxiv.org/abs/1901.03892v2>.
- [17] LI S, WANG J, LIU P, et al. Detection of multiple steganography methods in compressed speech based on code element embedding, Bi-LSTM and CNN with attention mechanisms [J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2021, 29: 1556-1569.
- [18] AGUSTSSON E, TIMOFTE R. Ntire 2017 challenge on single image super-resolution: dataset and study [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017: 126-135.
- [19] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context [C]//*Computer Vision-ECCV 2014*. Cham: Springer International Publishing, 2014: 740-755.